

Utilización del Método Científico en Proyectos y Experiencias de Innovación Educativa

Ángel Hernández García
Universidad Politécnica de Madrid
angel.hernandez@upm.es





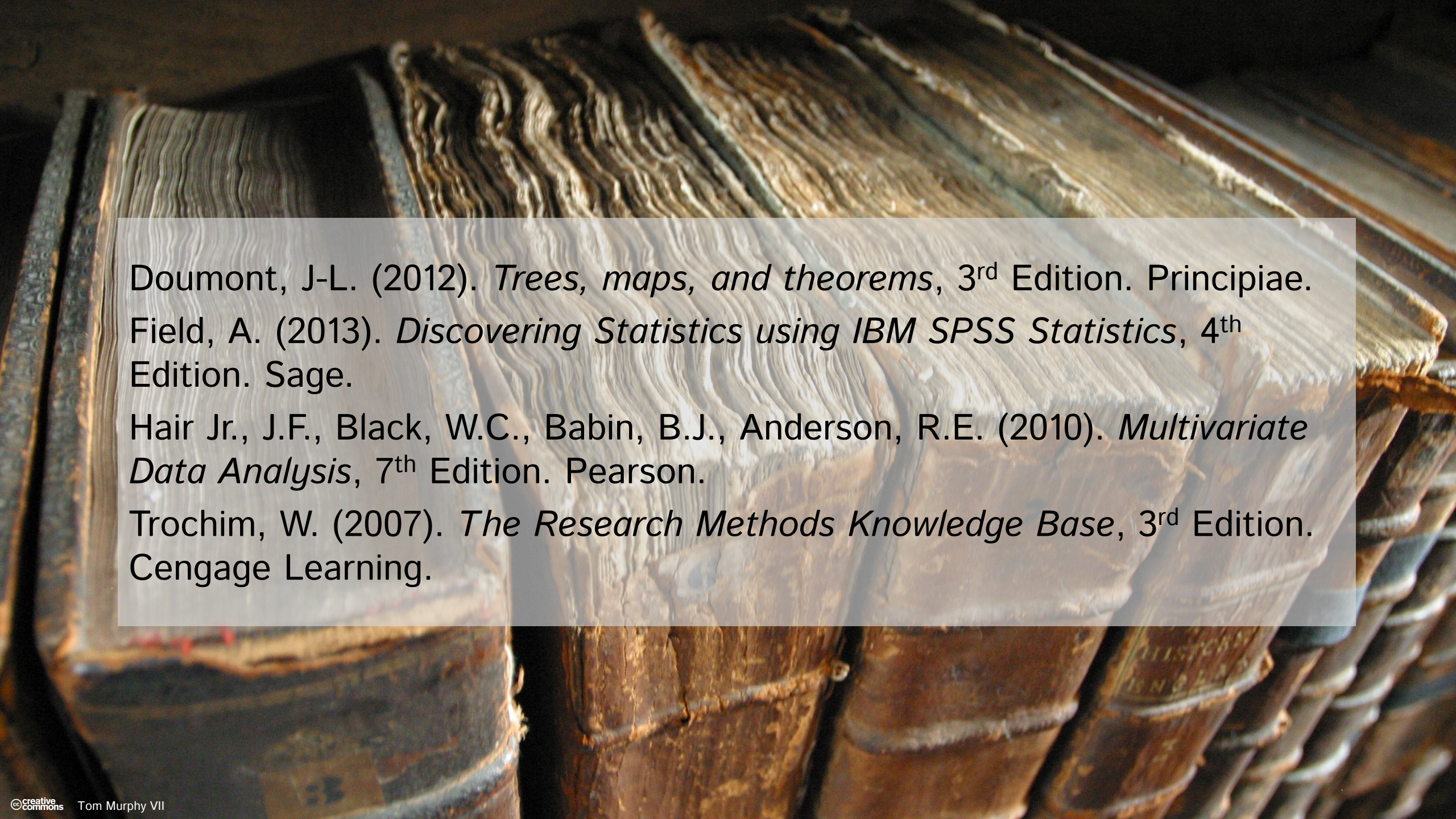
NOTA: Esta presentación fue creada originalmente y empleada para la impartición del seminario “Utilización del Método Científico en Proyectos y Experiencias de Innovación Educativa” en CINAIC 2015. La presentación tiene una licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Algunos de los contenidos gráficos pueden estar sujetos a derechos de autor. Gran parte del material es una adaptación personal del autor de la presente obra, bajo permiso del autor original, de *The Research Methods Knowledge Base*. Para más información al respecto puede visitar <http://www.socialresearchmethods.net/kb>. Si bien la licencia de esta presentación permite su uso y adaptación con fines no comerciales, siempre que se utilice la misma licencia de esta obra, se recomienda solicitar el permiso de uso de la *Knowledge Base* al autor original, William M.K. Trochim (https://cornell.qualtrics.com/SE/?SID=SV_e9T9VQ15cEXzOAY).

La forma apropiada de referenciar ambas obras es la siguiente:

Hernández-García, Á. (2015). Utilización del Método Científico en Proyectos y Experiencias de Innovación Educativa. *III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad (CINAIC 2015)*. Madrid, 14-16 de Octubre de 2015.

Trochim, William M. (2006). *The Research Methods Knowledge Base*, 2nd Edition. Disponible en línea: <http://www.socialresearchmethods.net/kb>.





Doumont, J-L. (2012). *Trees, maps, and theorems*, 3rd Edition. Principiae.

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*, 4th Edition. Sage.

Hair Jr., J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2010). *Multivariate Data Analysis*, 7th Edition. Pearson.

Trochim, W. (2007). *The Research Methods Knowledge Base*, 3rd Edition. Cengage Learning.





Conclusión

Análisis

Diseño

Medida

Muestra

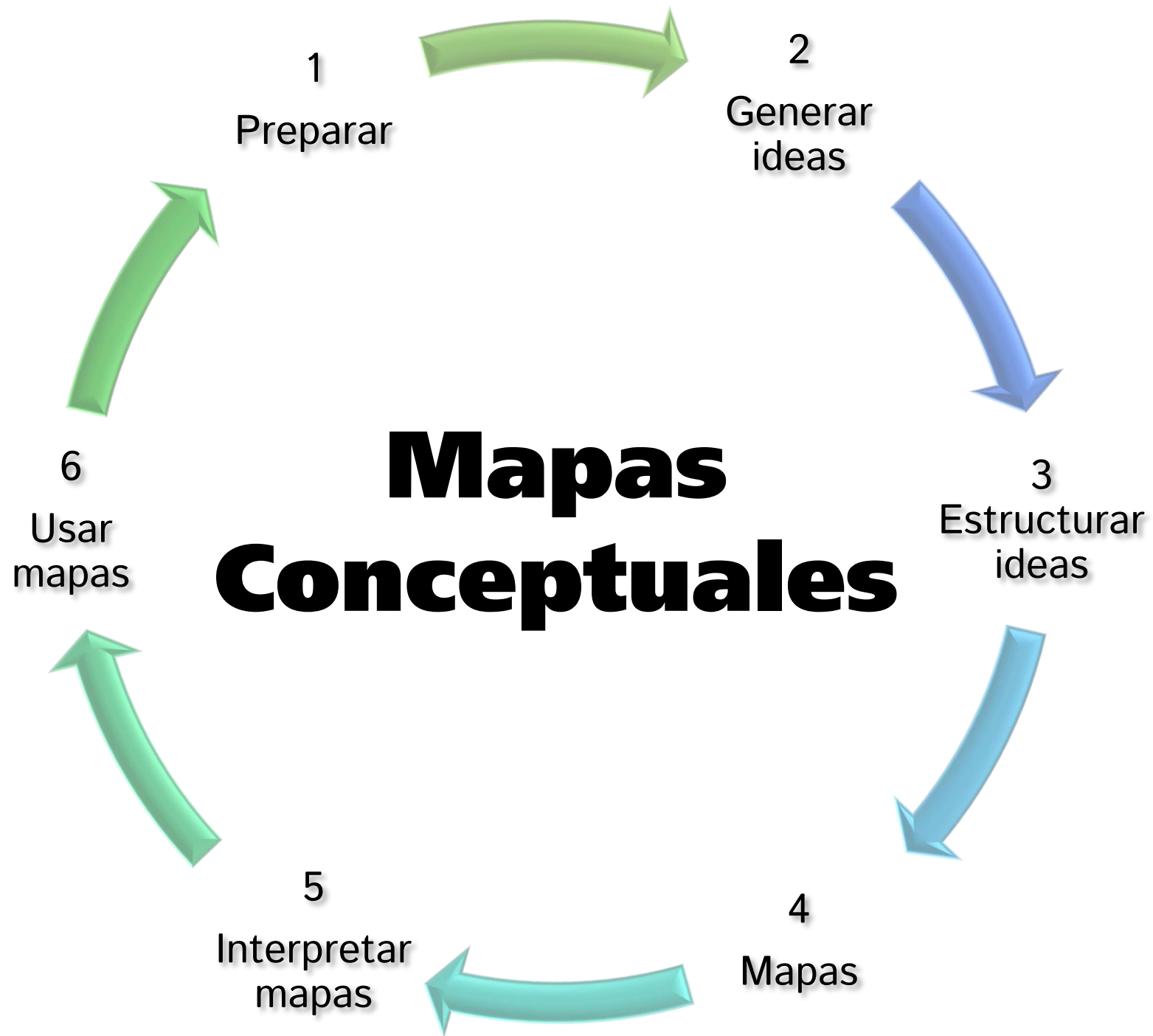
Problema

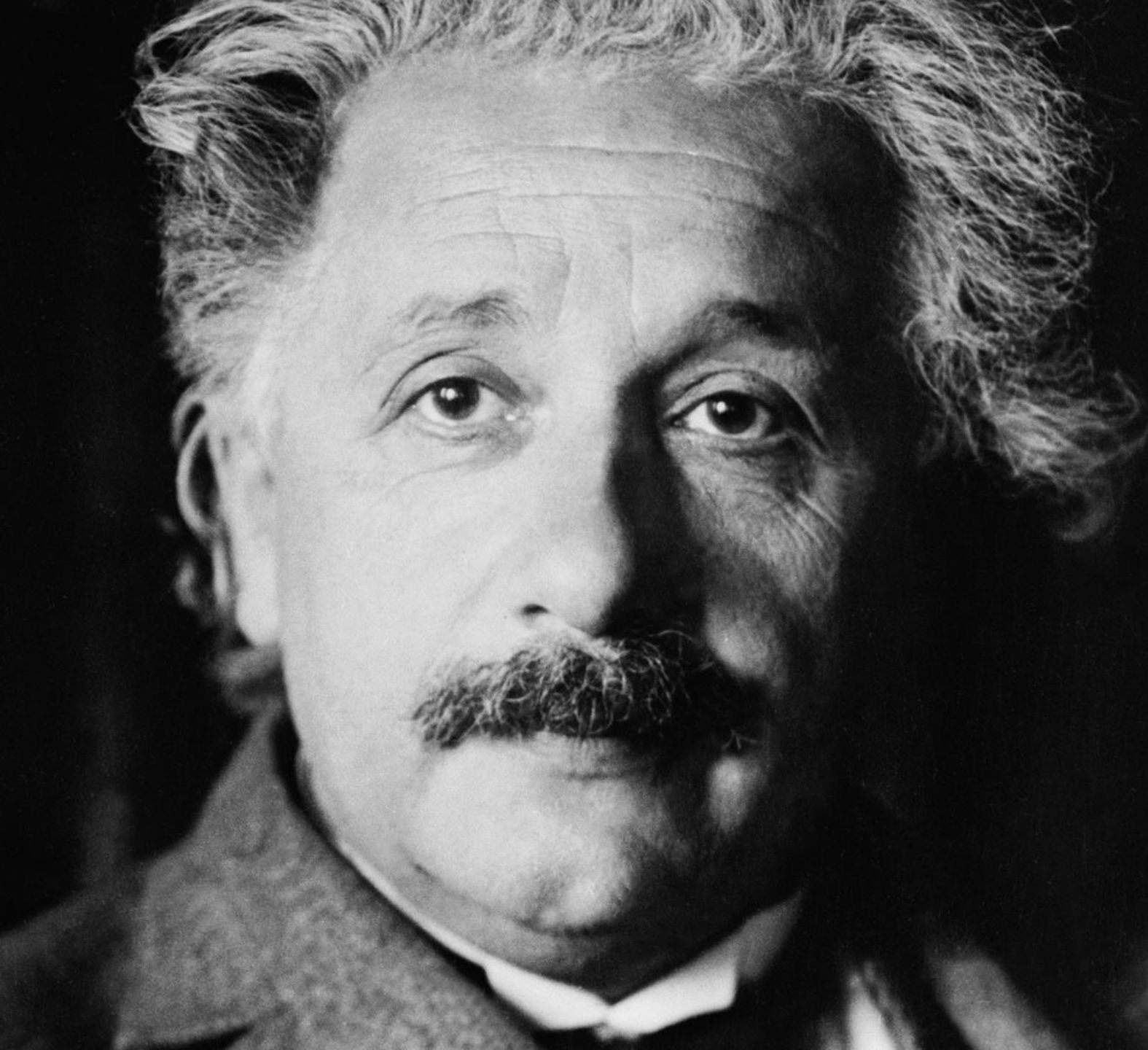


Fundamentos

El concepto



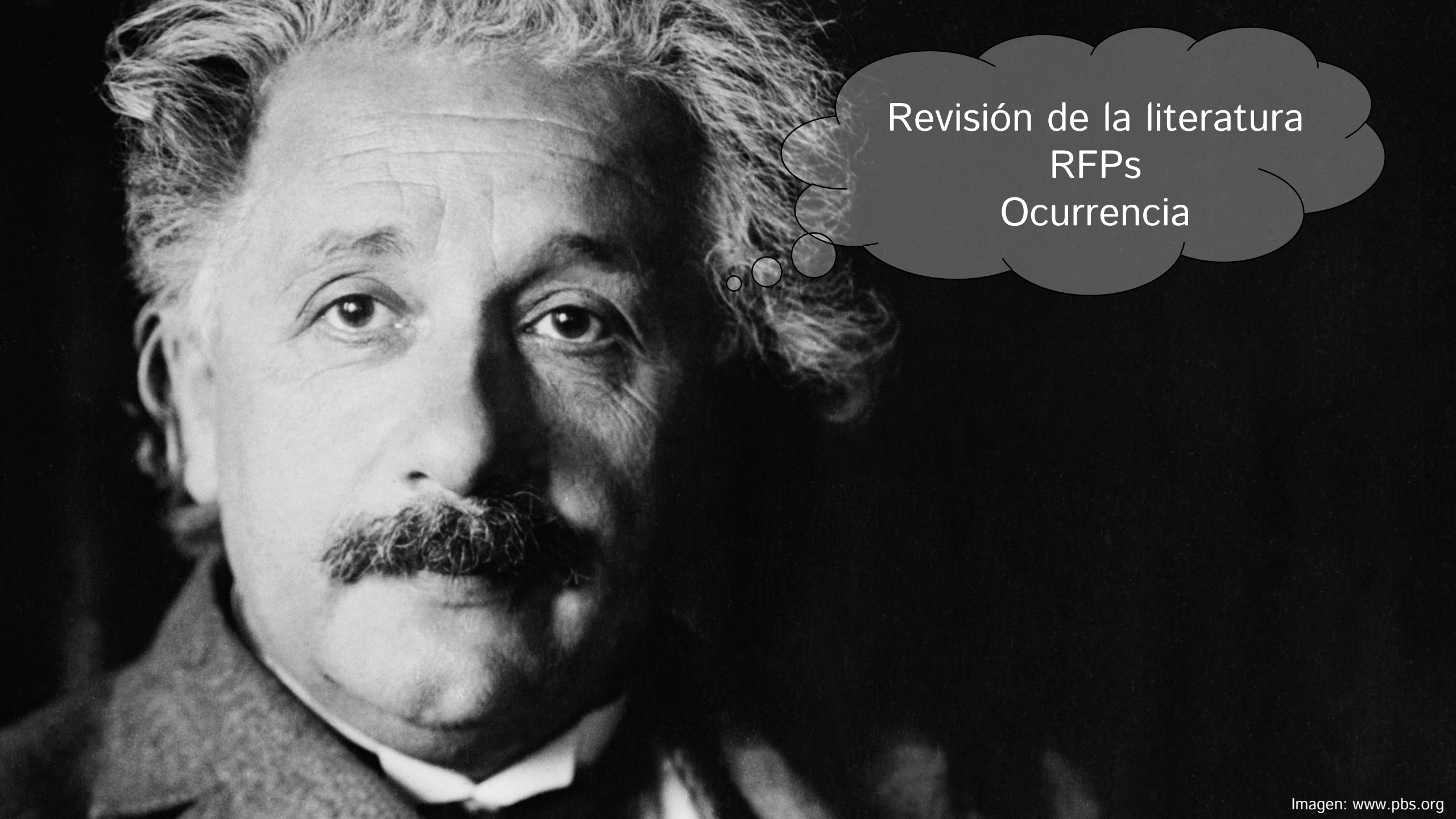




«La mera **formulación de un problema** es muchas veces más importante que su solución, que puede ser meramente una cuestión de habilidad matemática o experimental. Plantear **nuevas cuestiones, nuevas posibilidades, considerar viejos problemas** desde un nuevo ángulo, todo ello requiere de una imaginación creadora y marca los progresos reales de la ciencia.»



Viabile



Revisión de la literatura
RFPs
Ocurrencia

Revisión de la literatura

Tres tipos de preguntas

A group of diverse young children are sitting on the floor in a classroom, looking attentively towards the left. One girl in the foreground is raising her hand. The children are of various ethnicities and are dressed in casual clothing. The background is slightly blurred, showing more children and classroom furniture.

Descriptivas

Relacionales

Causales



Transversal

Longitudinal

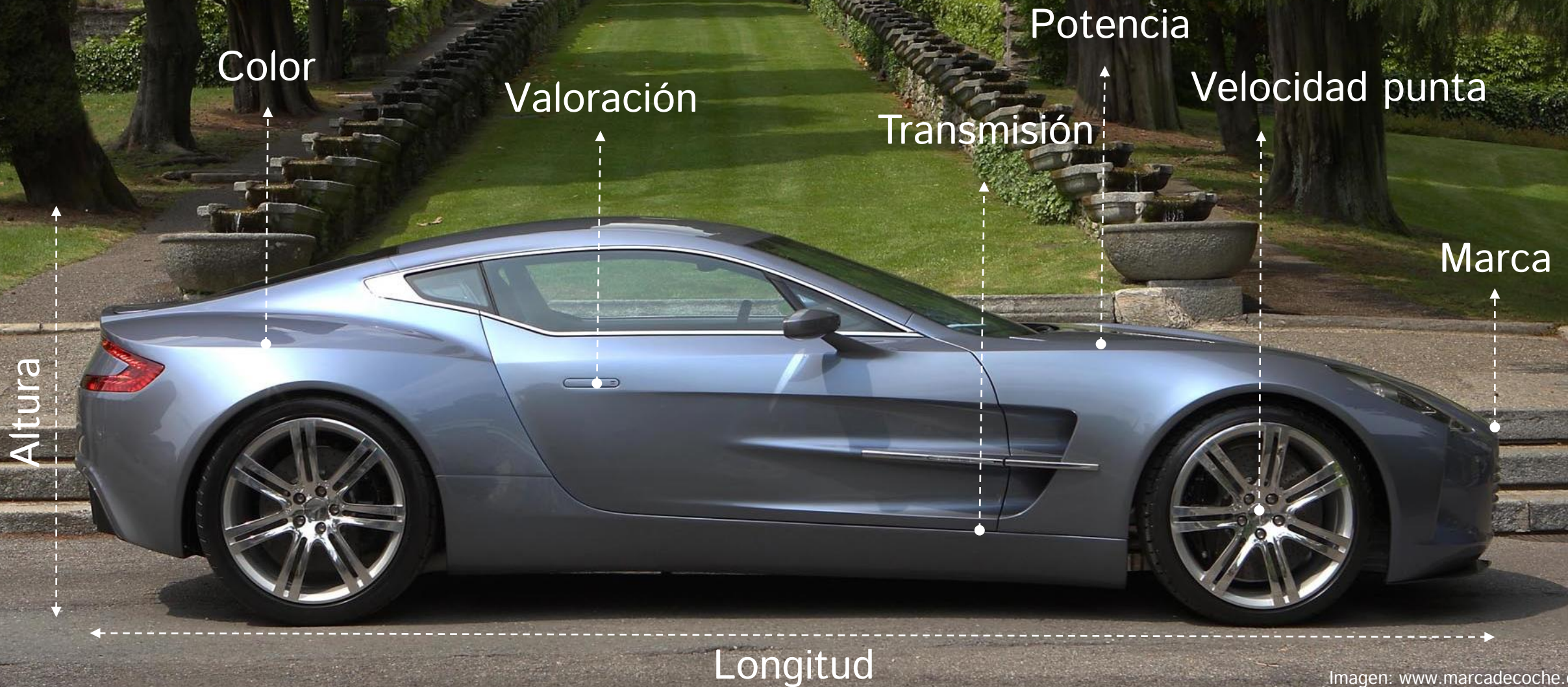
Medidas repetidas
Series temporales

A full-page background image of Michael Jordan in a red Chicago Bulls jersey, number 23, performing a slam dunk. He is in mid-air, with the basketball hoop and net visible above him. The background shows a large crowd of spectators in a stadium. A dashed white line runs vertically through the image, passing through Jordan's midline. A horizontal dashed white line is positioned above the hoop. In the bottom left corner, there is a circular inset image showing three men in suits, with one holding up a green Sprite sign with the number 10 on it.

Variable

Entidad que puede tomar diferentes valores

Variables y atributos



Dependiente

Manipulable
Natural

E
x
h
a
u
s
t
i
v
a

E
x
c
l
u
s
i
v
a

Independiente

Efecto
Resultado

Hipótesis

H_0 vs. H_1

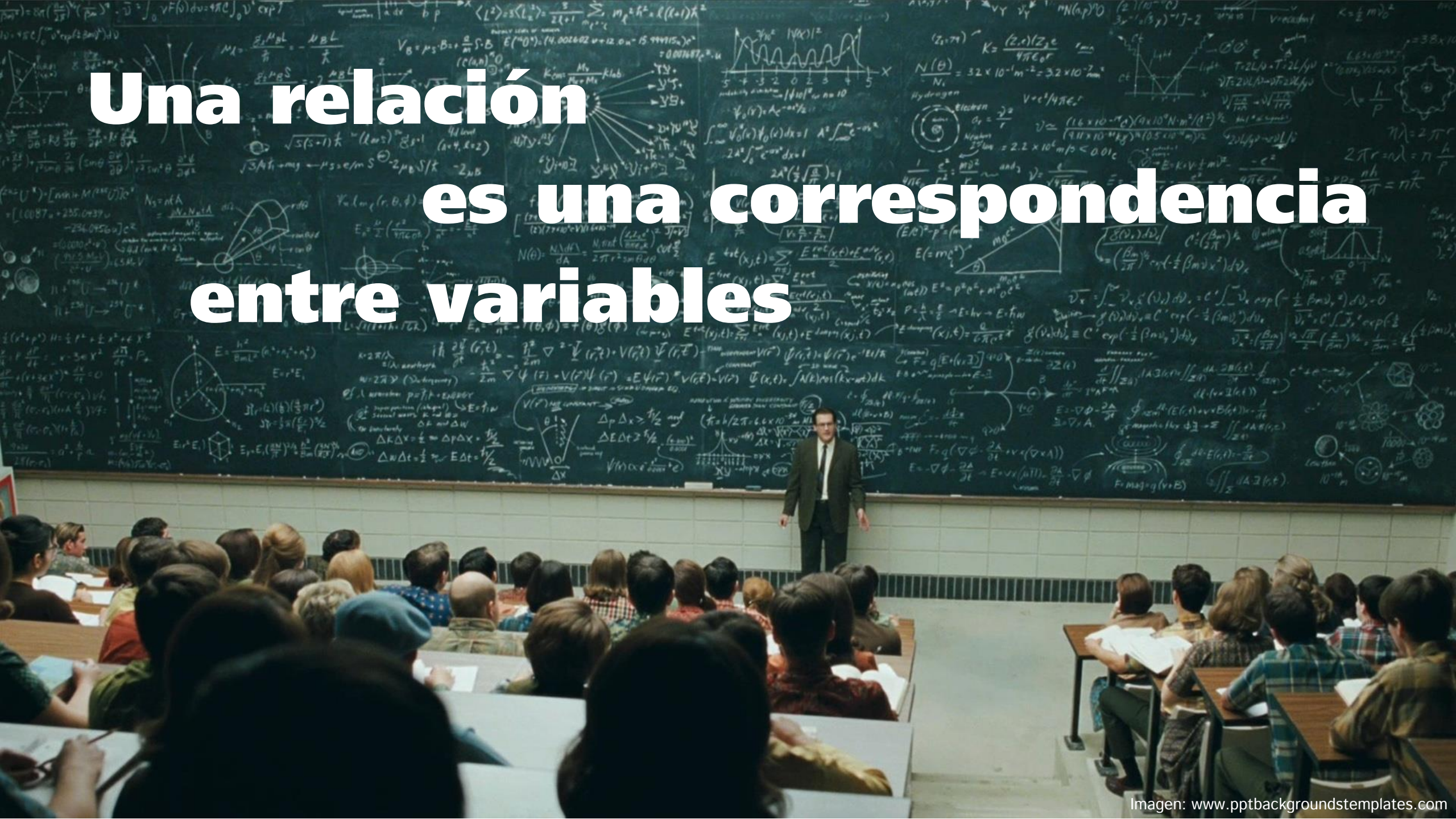
Excluyentes
Comprobables



VS



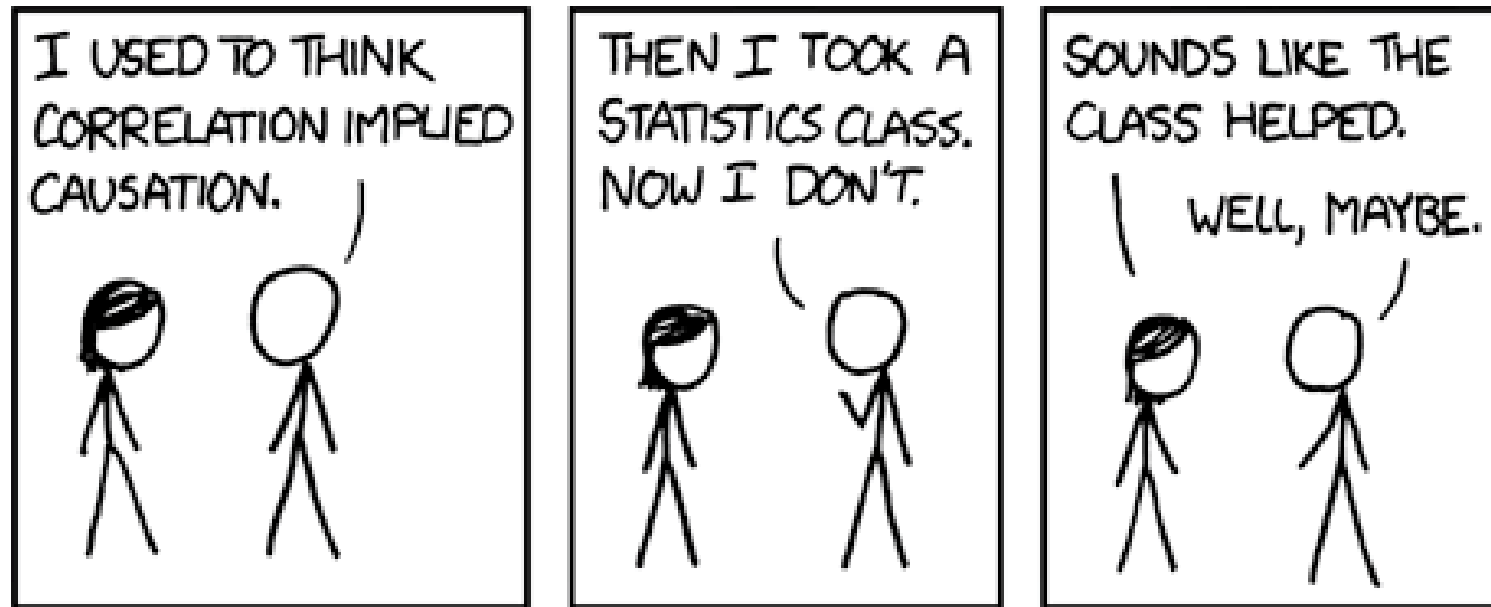
Una relación es una correspondencia entre variables



Correlacional
Causal
Asimétrica
Sin relación

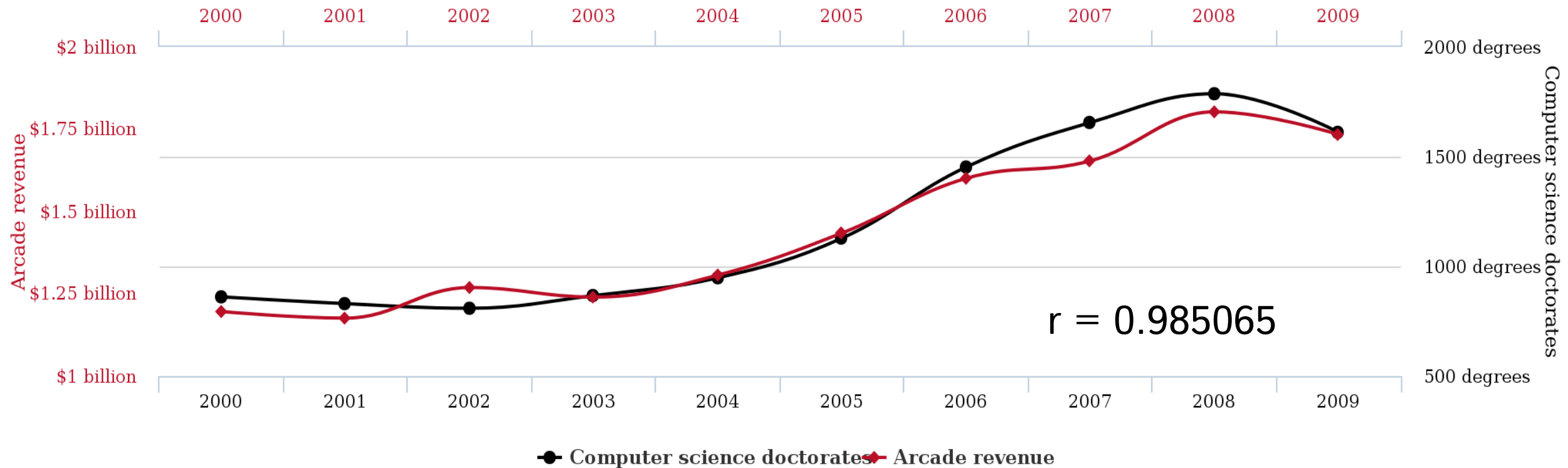


Correlación y causalidad



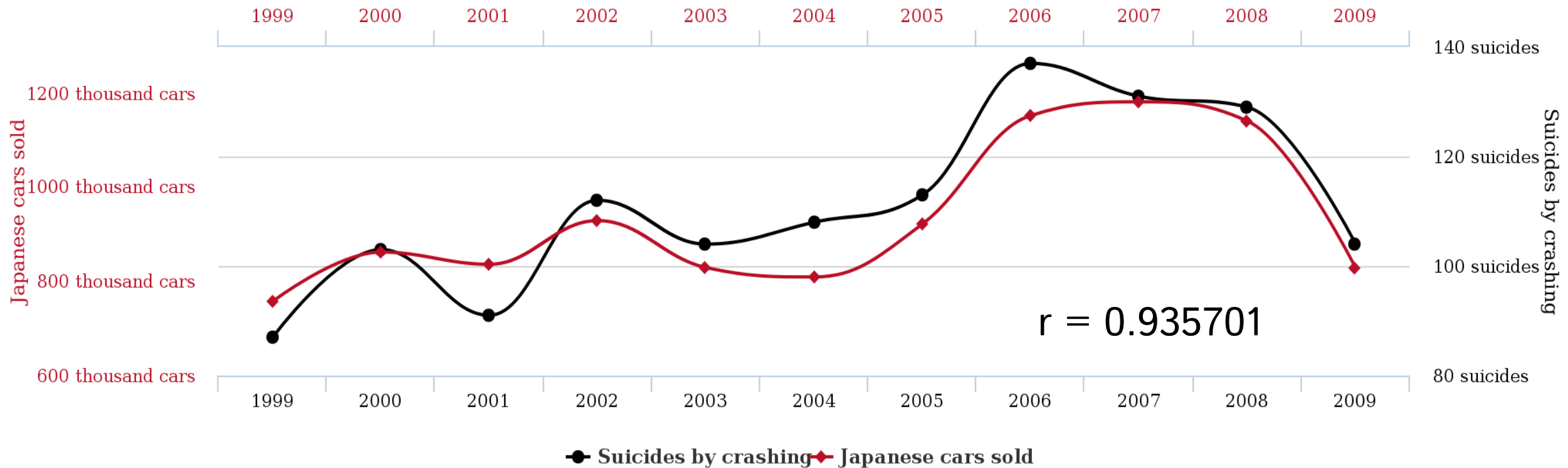
Correlación y causalidad

Total revenue generated by arcades
correlates with
Computer science doctorates awarded in the US



Correlación y causalidad

Japanese passenger cars sold in the US
correlates with
Suicides by crashing of motor vehicle



Datos cuantitativos en ciencias sociales: ¿son cualitativos? (y viceversa)

Juicios al elaborarlos

- Cómo definirlos
- Cómo distinguirlos
- Cómo construir indicadores
- Cómo asegurar que se entienden
- Contextos de uso
- Limitaciones culturales/idioma

Juicios al usarlos

- Cómo de bien mide el concepto
- Fiabilidad
- Apropiado para el contexto
- Apropiado para los participantes



Unidad de análisis

Individuos
Grupos
Artefactos
Unidades geográficas
Interacciones sociales
...

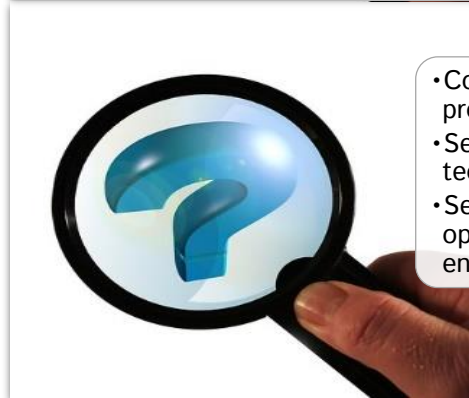
Componentes de un estudio de investigación

Problema de investigación



- Muy general
- Formulación

Pregunta de investigación



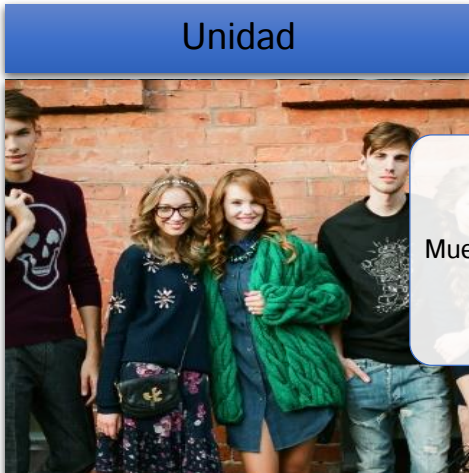
- Concreta el problema
- Se enmarca en teoría
- Se operacionaliza en hipótesis

Programa



- Representa la causa
- Constructos

Unidad



Muestreo

Resultados



Efectos

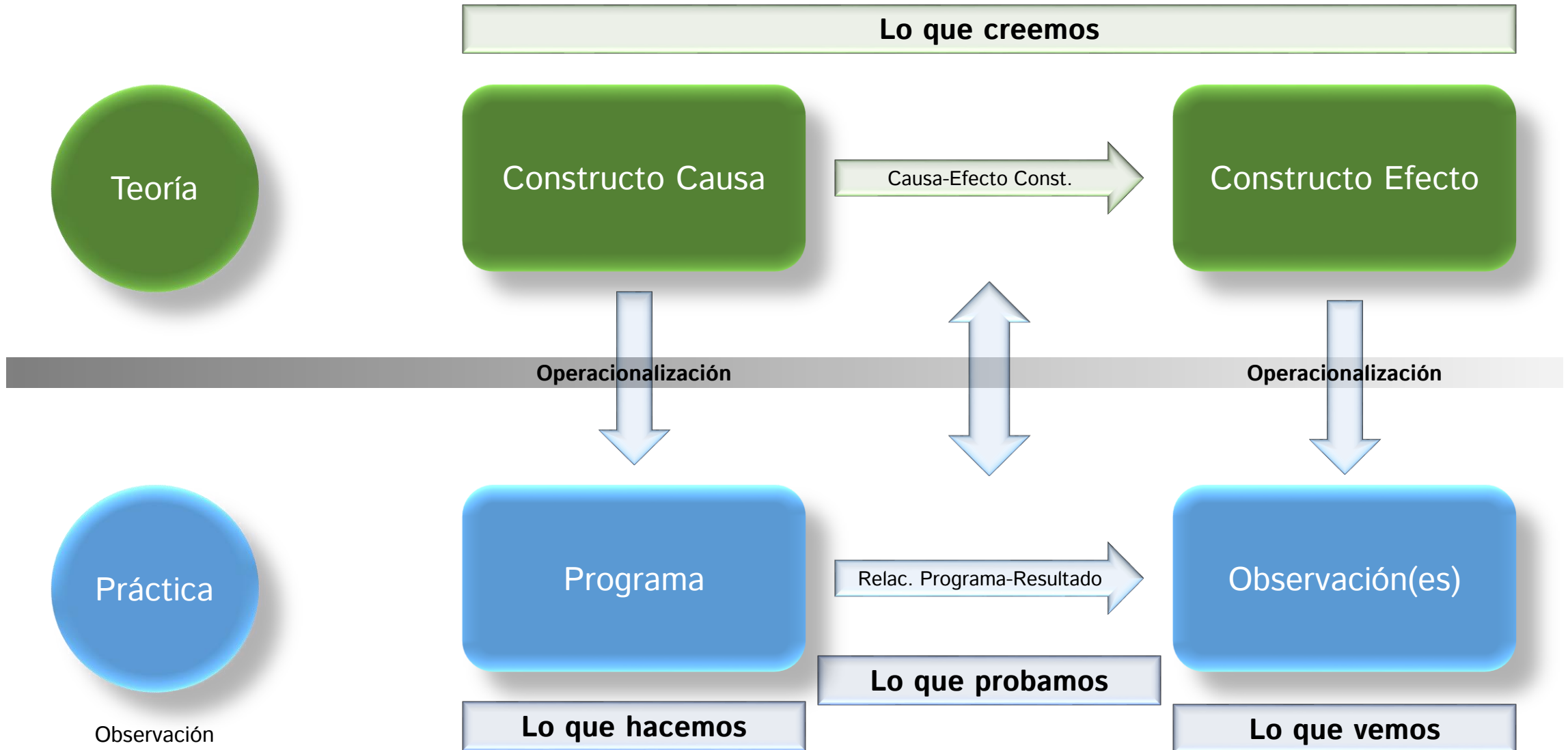
Diseño



Deducción vs. Inducción



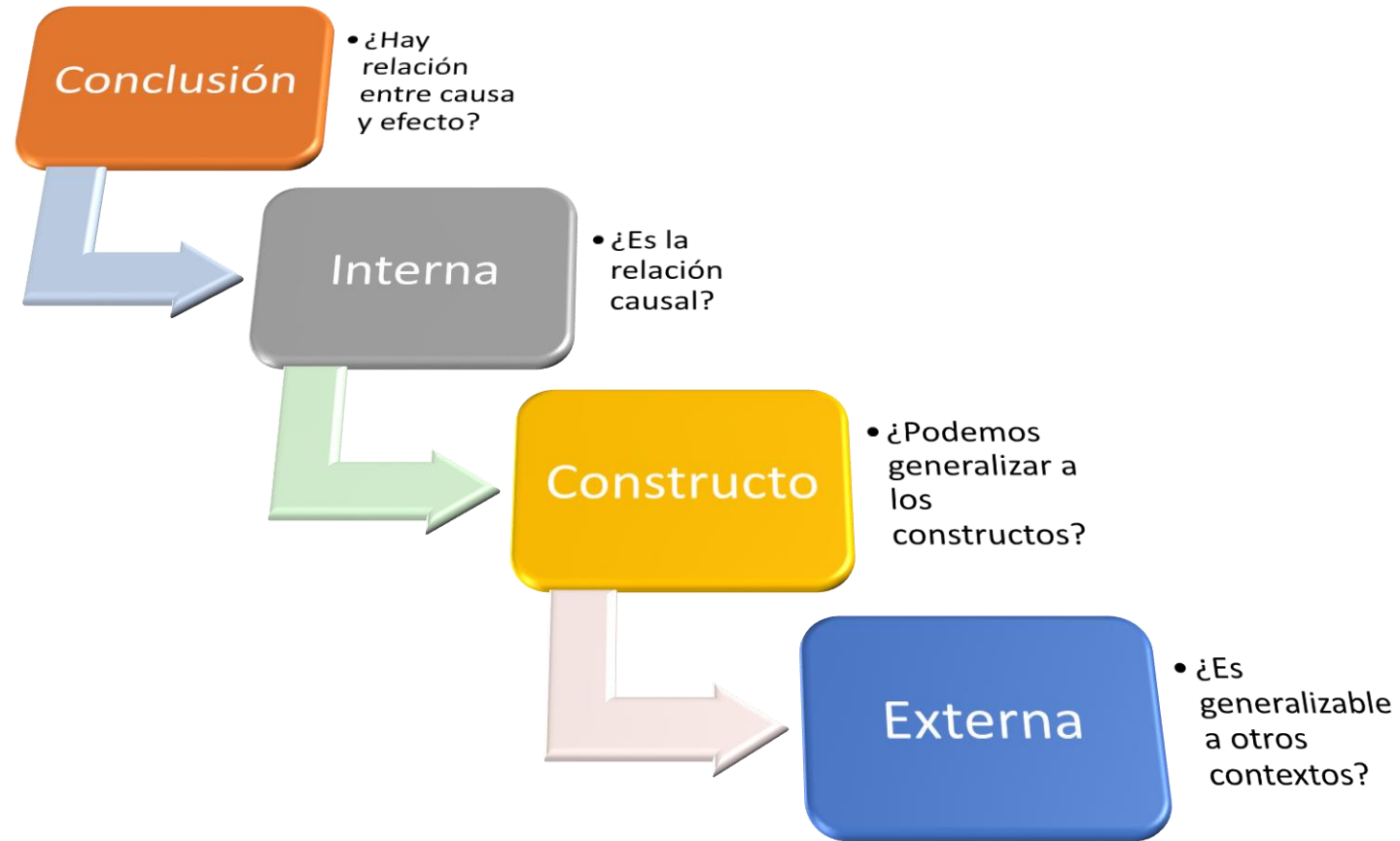
Validez



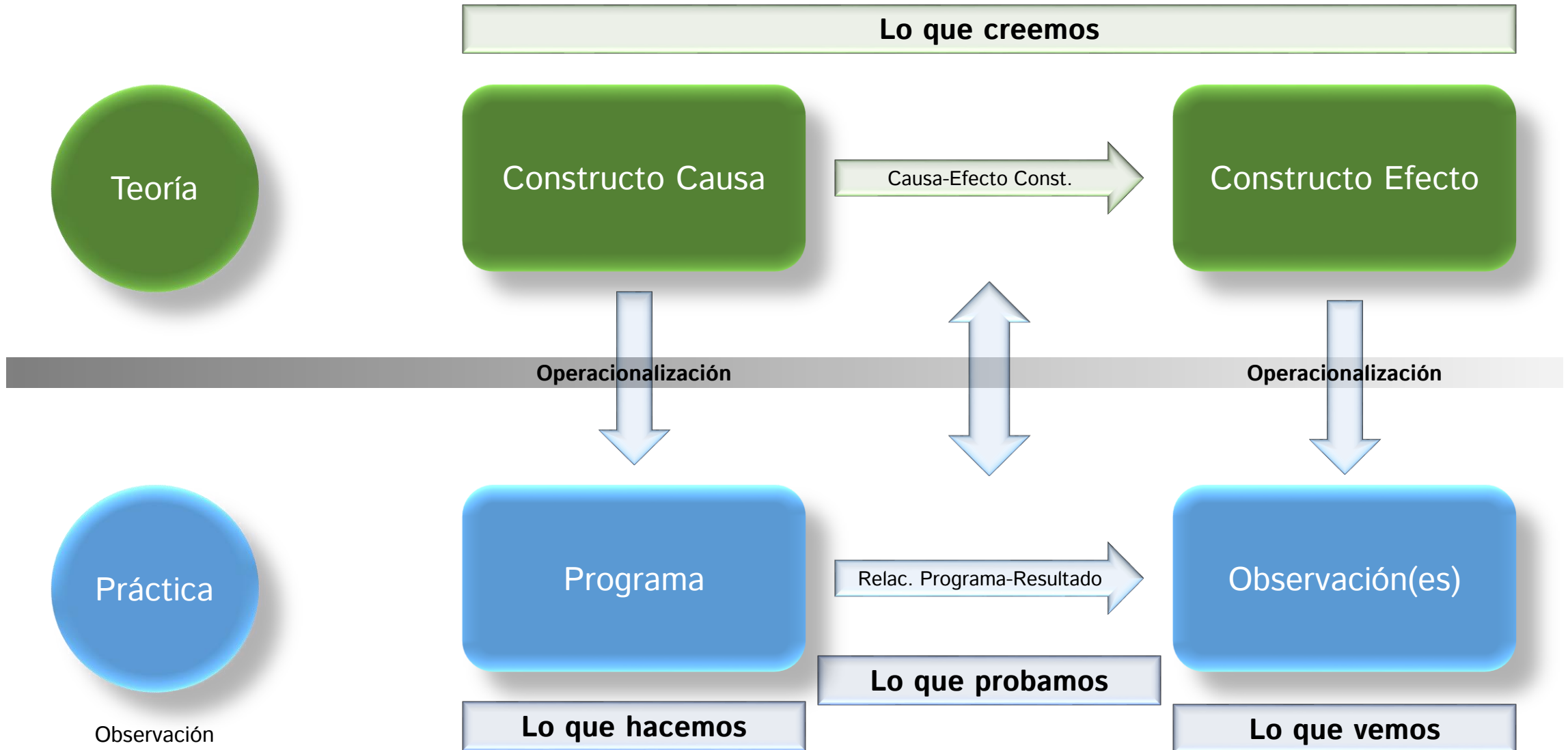
Validez

En este estudio...

Validez (tomar con cautela)



Validez





Adquisición y valoración **sistemática** de la información para proporcionar realimentación **útil** acerca de un objeto

Evaluación

Tipos de evaluación de investigación

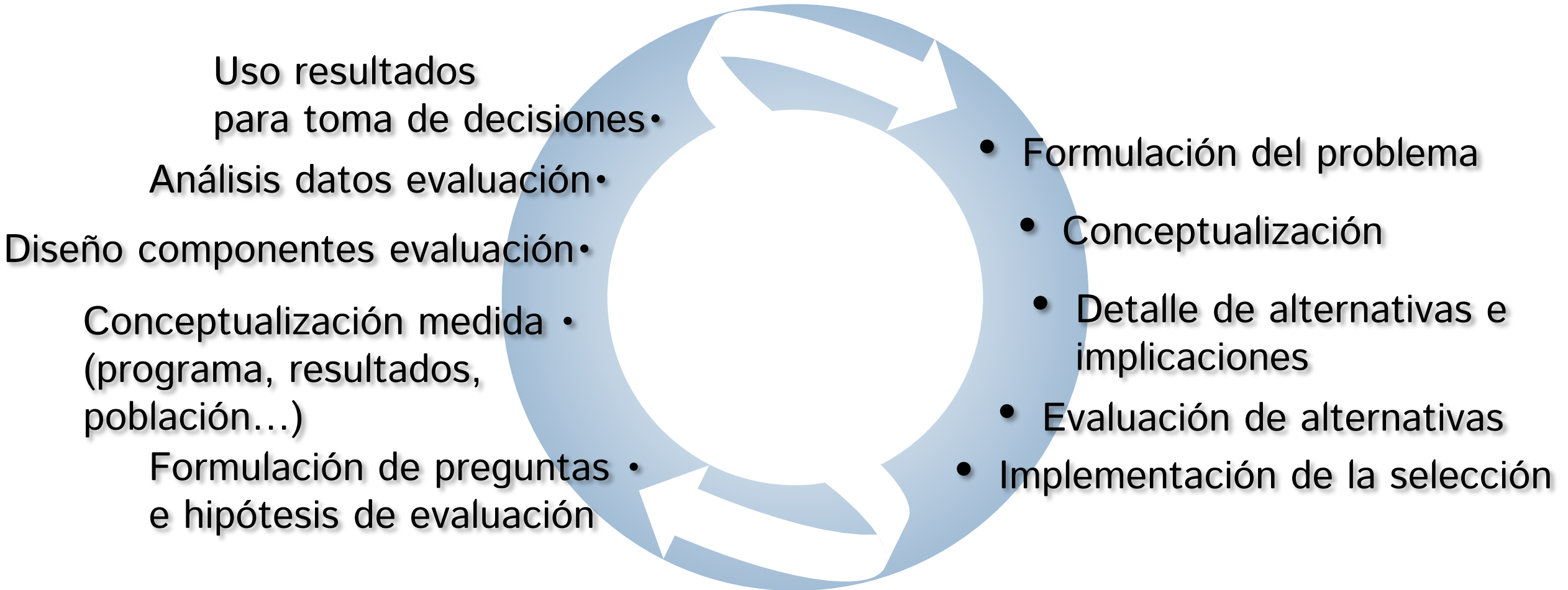
Formativa

Evaluación de necesidades
Conceptualización estructurada
Evaluación de implementación
Evaluación de proceso

Sumativa

Evaluación de los resultados
Evaluación de impacto (efecto neto)
Análisis coste-eficiencia
Análisis coste-beneficios
Análisis secundario
Meta-análisis

El ciclo planificación-evaluación





6'6"

6'6"

6'0"

6'0"

5'6"

5'6"

5'0"

5'0"

4'6"

4'6"

4'0"

4'0"

3'6"

3'6"

3'0"

3'0"

Validez externa

Modelado de la muestra

Identificación para **generalización**

Similaridad proximal

Contextos de generalización: tiempo, lugar, personas, contexto

Muestreo aleatorio* (pero de verdad) para mejorarla

*No confundir con asignación aleatoria



Conclusión

Análisis

Diseño

Medida

Muestra

Problema

Muestreo

A high-angle, wide shot of a large crowd at a concert. The crowd is dense, filling the lower two-thirds of the frame, with many people raising their hands in the air. In the background, a band is performing on a stage, illuminated by bright spotlights. The stage is elevated and has various musical equipment visible. The overall atmosphere is energetic and lively.

¿A quién queremos generalizar el estudio?

Población objetivo



Población objetivo



Población objetivo



Muestreo

A high-angle, wide shot of a large crowd at a concert. The crowd is dense and fills the lower two-thirds of the frame. Many people have their hands raised in the air, some holding up phones to record. The stage is at the top of the frame, illuminated by bright spotlights. A band is performing on stage, with several members visible. The venue appears to be a large hall or arena with ornate architectural details on the walls.

¿A quién podemos acceder?

Población del estudio



Muestreo

A high-angle, wide shot of a large crowd at a concert. The crowd is dense and fills the lower two-thirds of the frame. Many people have their arms raised in the air, some holding up phones to record. The stage is at the top of the frame, illuminated by bright spotlights. A band is performing on stage, with several members visible. The venue appears to be a large hall or arena with ornate architectural details on the walls.

¿A quién seleccionamos?

Muestra



Tipos de muestreo



Muestreo probabilístico: muestreo aleatorio

Muestreo aleatorio simple:

Excel (=ALEATORIO())

Muestreo aleatorio estratificado:

Muestreo simple con grupos homogéneos

Mayor representatividad

Muestreo aleatorio sistemático:

Orden aleatorio

Selección cada n puestos

Muestreo aleatorio de área (clúster)



Muestreo no probabilístico

En general, considerado menos riguroso

Accidental (muestreo de conveniencia)
Disponibilidad

Intencional
Bueno para muestra definida
Sobrerrepresentación

Instancia modal

Casos típicos

Muestreo de expertos (panel)

Muestreo de cuota: proporcional o no* proporcional

Muestreo de heterogeneidad

Muestreo bola de nieve





Conclusión

Análisis

Diseño

Medida

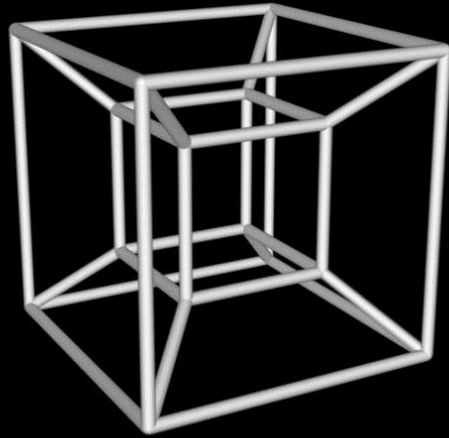
Muestra

Problema

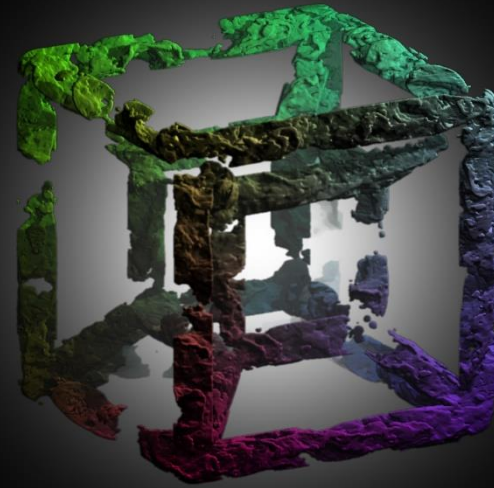
Medida



Validez del constructo



JFS
www.daviddarting.info



JFS
www.lashipire.polytechnique.fr

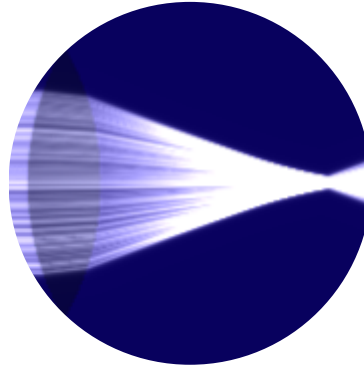
Tipos de validez

Traducción



“Pinta” y contenido

Convergente



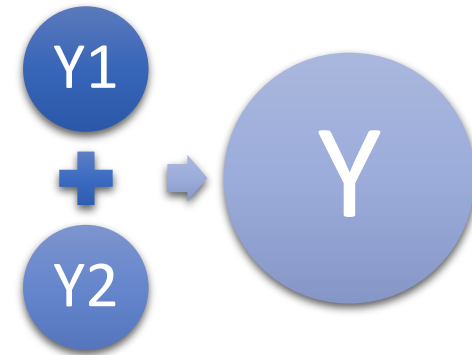
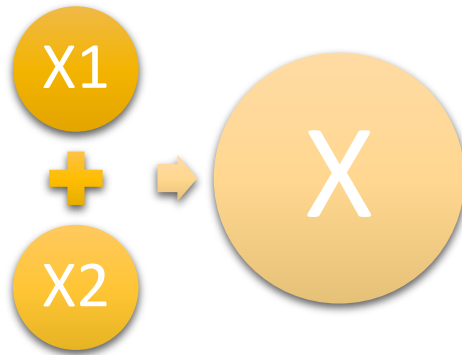
Alta correlación

Discriminante



Baja correlación

Validez convergente y discriminante



	X1	X2	Y1	Y2
X1	1,000	0,931	0,121	0,051
X2	0,931	1,000	0,090	0,106
Y1	0,121	0,090	1,000	0,955
Y2	0,051	0,106	0,955	1,000

Amenazas a la validez del constructo

A person in a full-body protective suit, including a helmet and goggles, is running through a field of debris and smoke. The person is wearing a dark, heavy suit with a large, clear visor. They are running towards the viewer, with their right leg forward. The background is filled with a large, billowing cloud of white and grey smoke or dust, and numerous small, dark fragments of debris are suspended in the air around the person. The ground is covered in a layer of brown, sandy material, possibly dirt or ash. The overall scene suggests a hazardous environment, such as a chemical spill or a battlefield.

Preoperacionalización
Explicación

Sesgo mono-operación
Sesgo mono-método

Efectos de interacción
entre tratamientos

Amenazas sociales
“Adivinar hipótesis”
Aprensión a la evaluación
Deseabilidad social
Método común

Fiabilidad del instrumento de medida



Teoría del valor verdadero

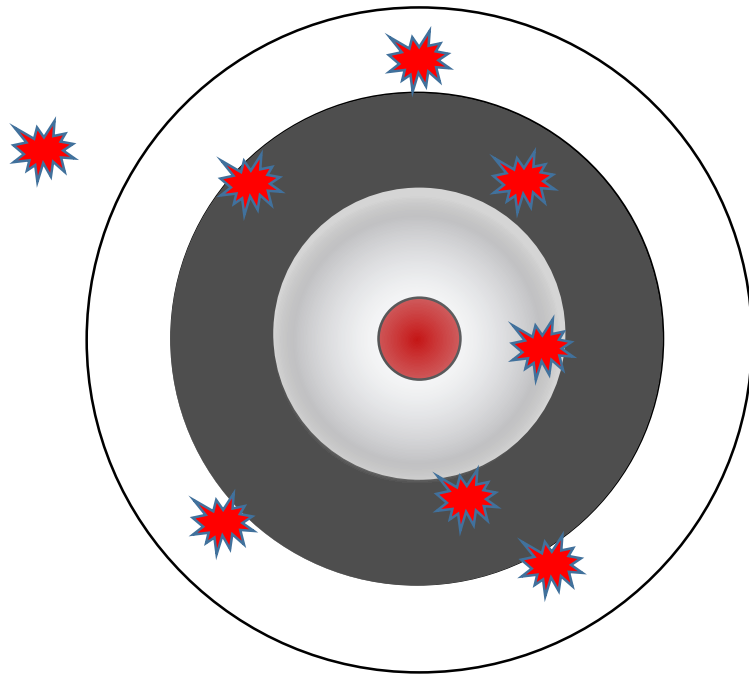


$$\text{Valor observado} = \text{Valor verdadero} + \text{Error}$$

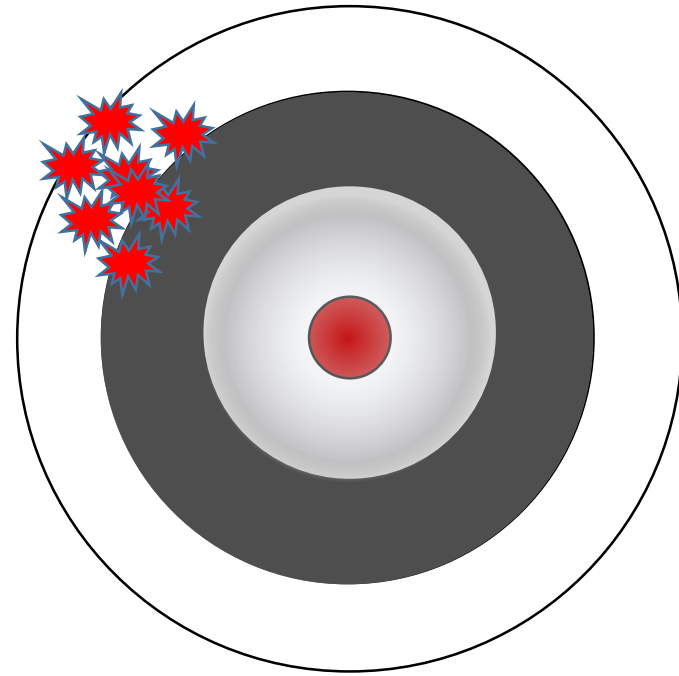
(medible) (estimable)

Error

Aleatorio



Sistemático



**La fiabilidad
es consistencia,
reproducibilidad**



Tipos de fiabilidad

Inter-observador (grado de acuerdo)

Test-retest

Formularios paralelos

Consistencia interna

Correlación inter-ítem media

Correlación ítem-total media

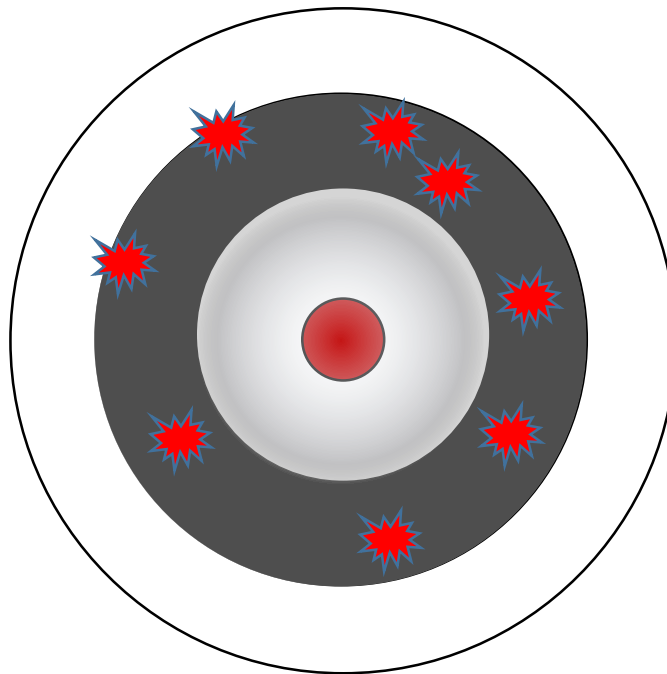
Split-half

Alfa de Cronbach

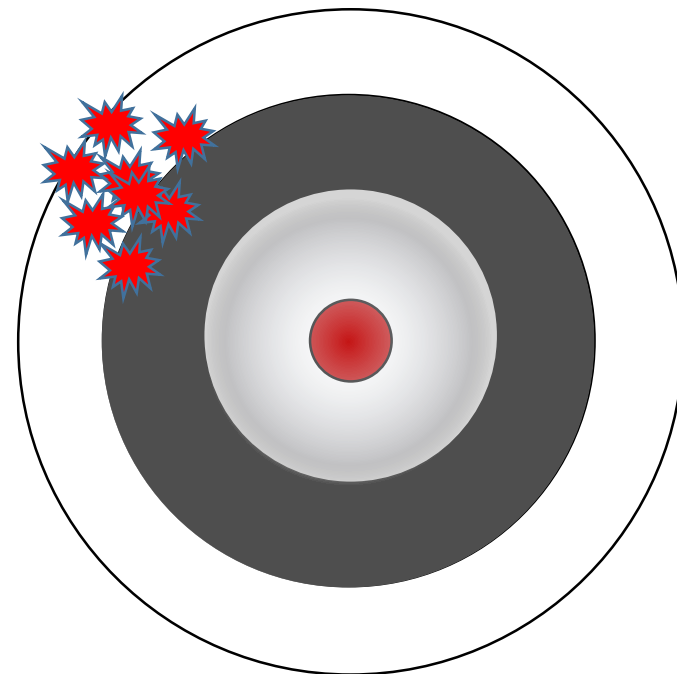


Validez y fiabilidad

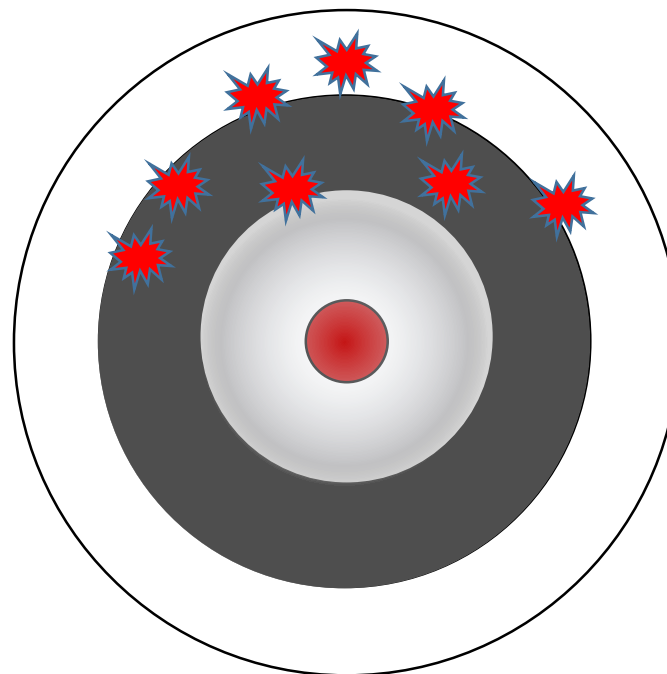
VnF



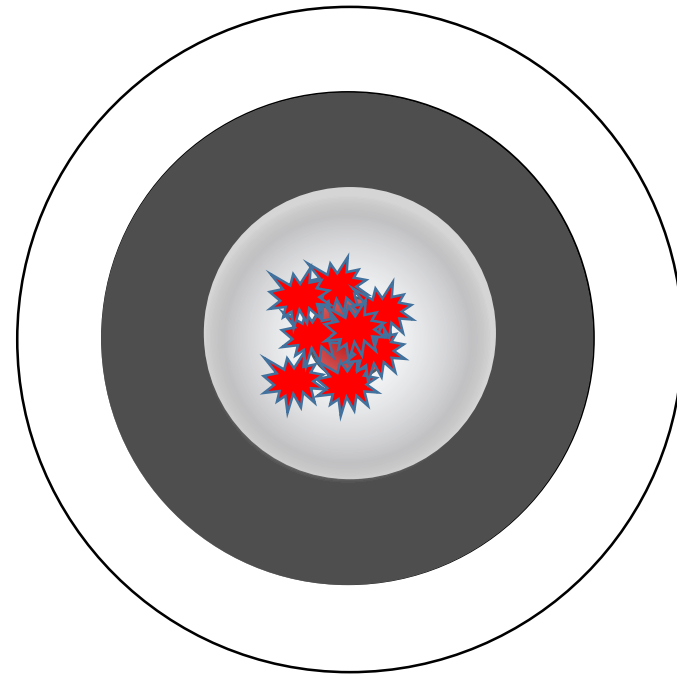
nVF



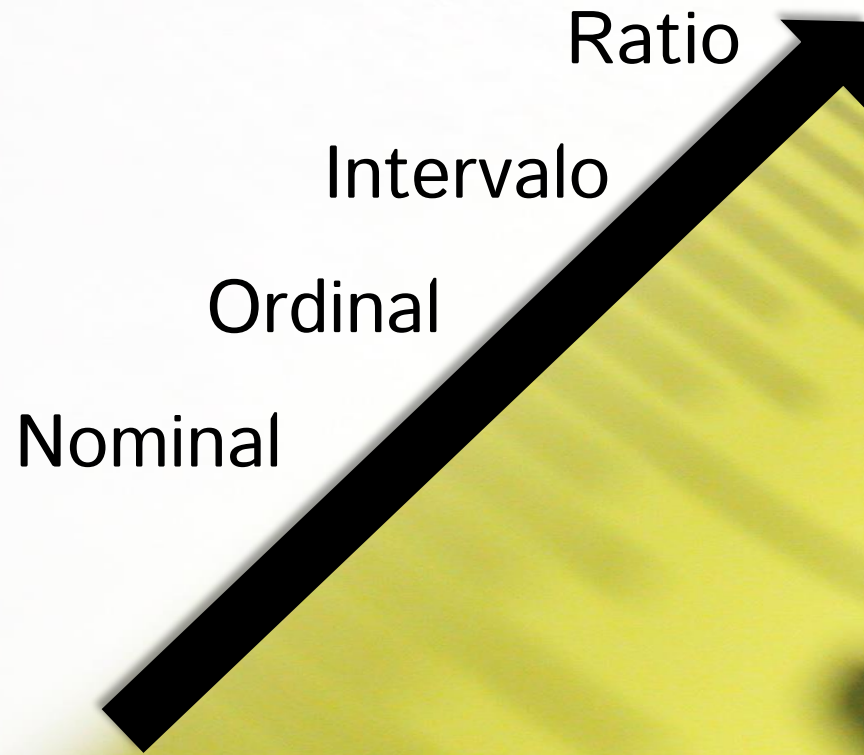
nVnF



VF



Niveles de medida: Escalas





Tipos de encuestas: Cuestionarios y entrevistas

Questionarios y entrevistas: ventajas y desventajas

Issue	Questionnaire			Interview	
	Group	Mail	Drop-Off	Personal	Phone
Are Visual Presentations Possible?	Yes	Yes	Yes	Yes	No
Are Long Response Categories Possible?	Yes	Yes	Yes	???	No
Is Privacy A Feature?	No	Yes	No	Yes	???
Is the Method Flexible?	No	No	No	Yes	Yes
Are Open-ended Questions Feasible?	No	No	No	Yes	Yes
Is Reading & Writing Needed?	???	Yes	Yes	No	No
Can You Judge Quality of Response?	Yes	No	???	Yes	???
Are High Response Rates Likely?	Yes	No	Yes	Yes	No
Can You Explain Study in Person?	Yes	No	Yes	Yes	???
Is It Low Cost?	Yes	Yes	No	No	No
Are Staff & Facilities Needs Low?	Yes	Yes	No	No	No
Does It Give Access to Dispersed Samples?	No	Yes	No	No	No
Does Respondent Have Time to Formulate Answers?	No	Yes	Yes	No	No
Is There Personal Contact?	Yes	No	Yes	Yes	No
Is A Long Survey Feasible?	No	No	No	Yes	No
Is There Quick Turnaround?	No	Yes	No	No	Yes

Aspectos a tener en cuenta



Población

- Enumerable (“encontrable”)
- Alfabetización
- Idioma
- Cooperación
- Restricciones geográficas

Preguntas

- Tipo
- Complejidad
- Filtrado (requisitos)
- Secuenciación
- Longitud
- Respuestas largas

Muestra

- Accesibilidad
- Quién responde
- Tasas de respuesta

Contenido

- Conocimientos previos
- Consulta registros

Sesgos

- Método
- Deseabilidad
- Falsos participantes

Administrativos

- Costes
- Tiempo
- Instalaciones
- Personal disponible

Cómo crear una encuesta



1. Determinar el contenido, alcance y propósito de las preguntas.
2. Elegir el formato de respuestas.
3. Formulación de las preguntas.

Tipos de preguntas



Dicotómicas.

Niveles de medida: nominal, ordinal (preferencias), intervalo (Likert), diferencial semántico, Guttman (acumulativa).

Preguntas filtro (condicionales).

No más de dos niveles.

Gráfico de salto.

Escalas

Asignación de objetos a números,
de acuerdo a una regla

Escalas

Thurstone

Likert

Pretest

Guttman

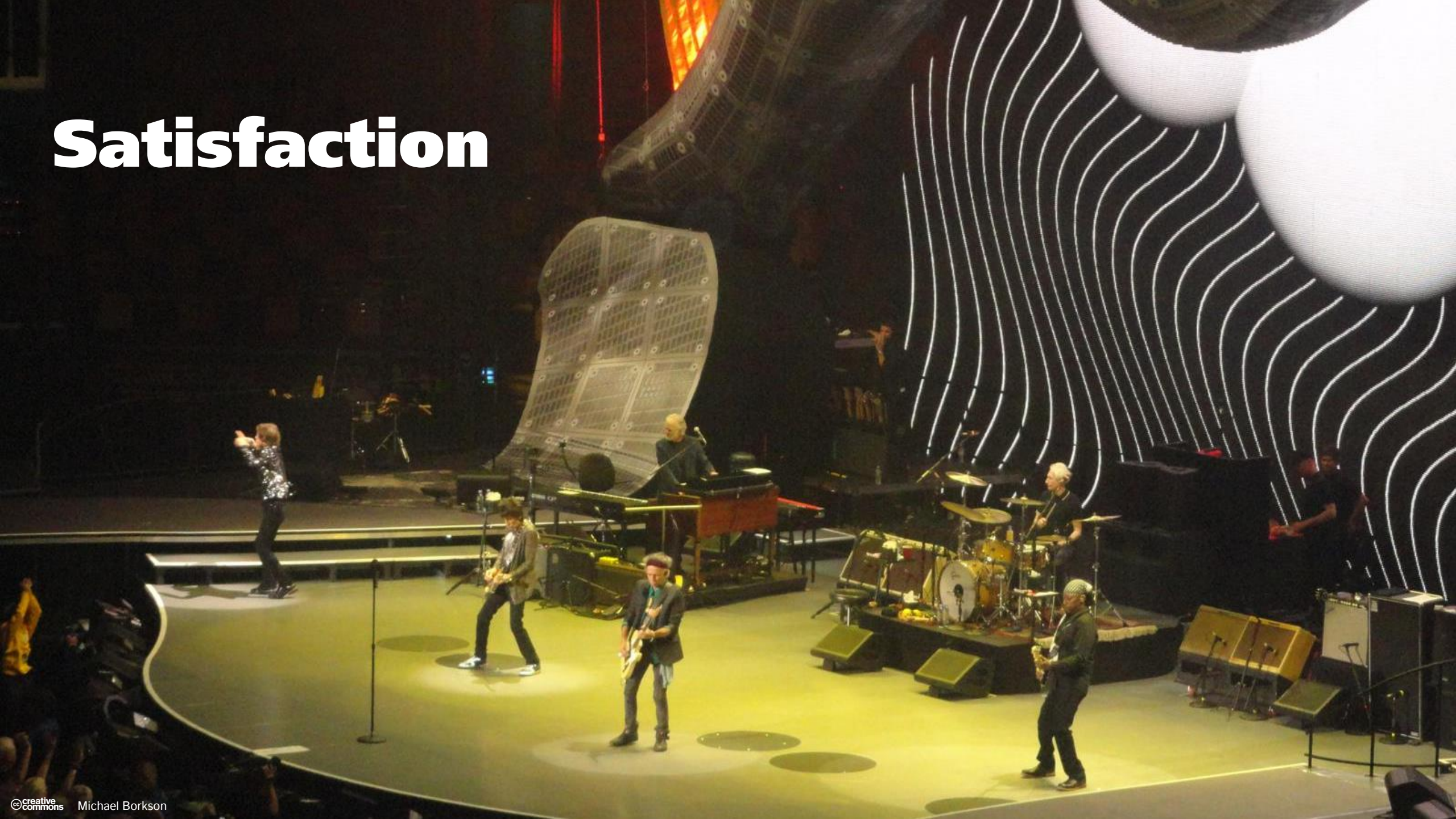


Escalas

Dimensionalidad



Satisfaction





Conclusión

Análisis

Diseño

Medida

Muestra

Problema



Diseño de la investigación

Causa y efecto

Validez interna

Aproximación a
la veracidad
acerca de inferencias
en relaciones causales

Precedencia temporal

Covariación causa-
efecto

Sin explicaciones
alternativas

Grupos de control



Grupo simple vs. múltiples grupos

Posttest y pretest-posttest

Amenazas grupo simple

Historia

Maduración

Testeo (sólo pre-post)

Instrumentación (sólo pre-post)

Mortalidad

Regresión (a la media)

Grupos de control (pre-post)

Amenazas múltiples grupos

Selección-Historia

Selección-Maduración

Selección-Testeo (sólo pre-post)

Selección-Instrumentación (sólo pre-post)

Selección-Mortalidad

Selección-Regresión (a la media)



Múltiples grupos

Los grupos deben ser comparables

Experimento aleatorio (verdadero)

Diseños cuasi-experimentales



Diseño de investigación

Observaciones/medidas (O)

Tratamiento/programa (X)

[Grupos]

[Asignación a grupos]

Aleatoria (R)

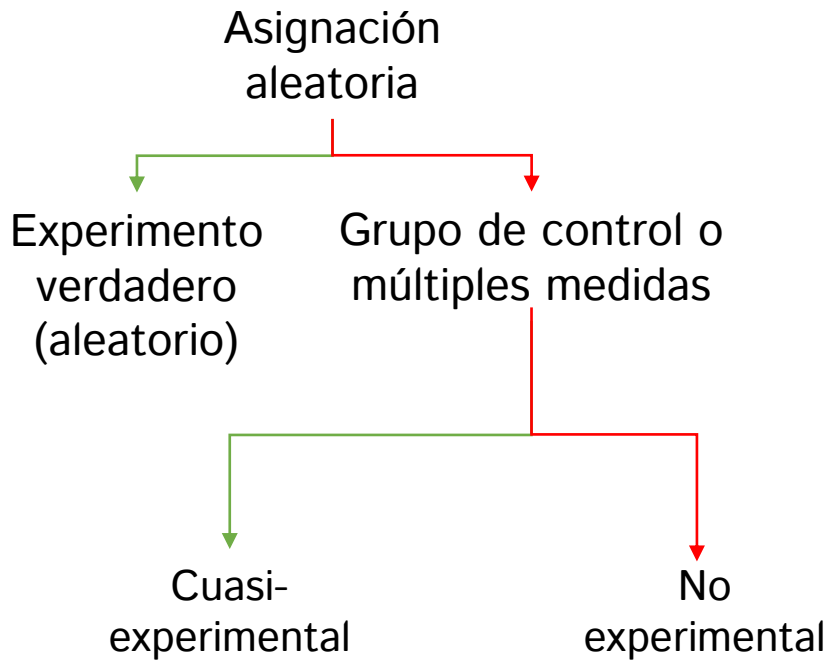
Grupos no equivalentes (N)

Corte (C)

Ejemplo: pre-post

R	O	X	O
R	O		O

Tipos de diseño



R	X	O
R		O

Diseño experimental aleatorizado (sólo posttest)

R	O	X	O
R	O		O

Diseño experimental aleatorizado (pretest-posttest)

N	O	X	O
N	O		O

Diseño cuasi-experimental
Grupos no equivalentes
(pretest-posttest)

X	O
---	---

No experimental
(sólo posttest)

Diseño experimental

Mayor robustez
validez interna

Si X, entonces Y
Si no X, entonces no Y

La asignación aleatoria
es clave

Equivalencia probabilística

Equilibrio validez interna
vs. validez externa

Diseño experimental: Dos grupos (básico)

R	X	O
R		O

Grupos equivalentes

Estamos interesados en determinar si los dos grupos son diferentes tras el programa (tratamiento).

T-test (o equivalente)
ANOVA
Regresión

Amenazas:
Selección-Mortalidad (tasas abandono)

Diseños experimentales

El objetivo es
maximizar la SNR



Diseños factoriales

Maximizar la señal

Diseños factoriales

Ejemplo:

Efecto del tiempo de estudio y del lugar:

R	X_{11}	0
R	X_{12}	0
R	X_{21}	0
R	X_{22}	0

Posibles resultados...

Diseños factoriales

		Tiempo		
		1	2	
Lugar	1	5	5	5
	2	5	5	5
		5	5	

Sin efectos

		Tiempo		
		1	2	
Lugar	1	5	5	5
	2	5	7	6
		5	6	

		Tiempo		
		1	2	
Lugar	1	7	5	6
	2	5	7	6
		6	6	

Efectos de interacción

		Tiempo		
		1	2	
Lugar	1	5	7	6
	2	5	7	6
		5	7	

		Tiempo		
		1	2	
Lugar	1	5	5	5
	2	7	7	7
		6	6	

		Tiempo		
		1	2	
Lugar	1	5	7	6
	2	7	9	8
		6	8	

Efectos principales

Diseño en bloques aleatorizados y diseños de covarianzas

Minimizar el ruido

Bloques aleatorizados

Equivalente a muestreo
aleatorio estratificado

Homogéneos

R X O

R X O

R X O

R X O

Heterogéneos

Diseños de covarianzas

ANCOVA

R	O	X	O
R	O		O

Permite eliminar heterogeneidad causada por covariables.
Ejemplo: ajuste del posttest para variabilidad en pretest.

Diseños cuasi-experimentales

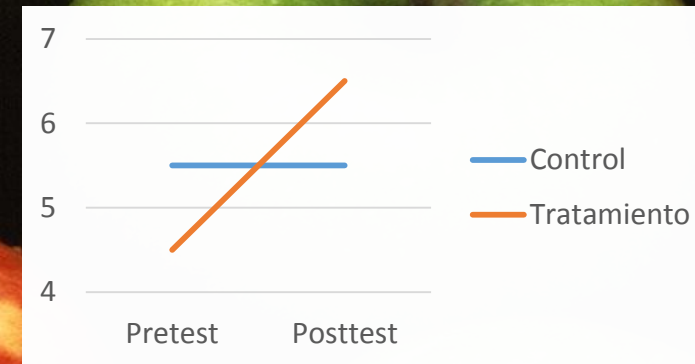
¿Inferiores?

Diseño de grupos no equivalentes

N	O	X	O
N	O		O

Los grupos deben ser tan similares como sea posible...
... pero no podemos estar seguros de que sean comparables

Diseño de grupos no equivalentes



Amenazas:
Selección

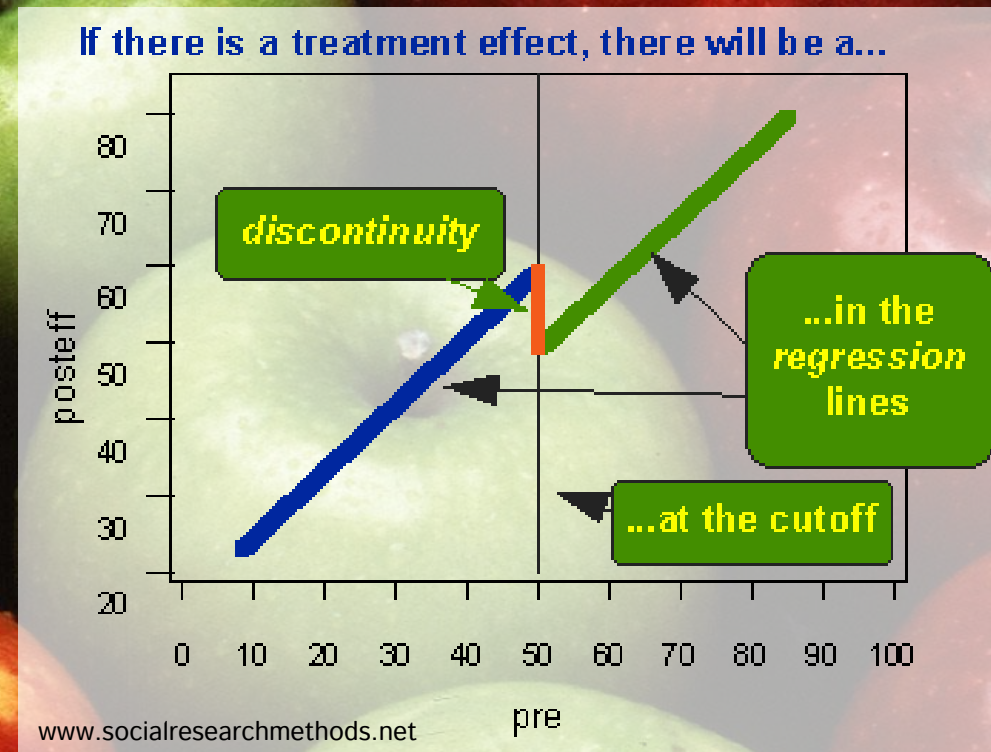
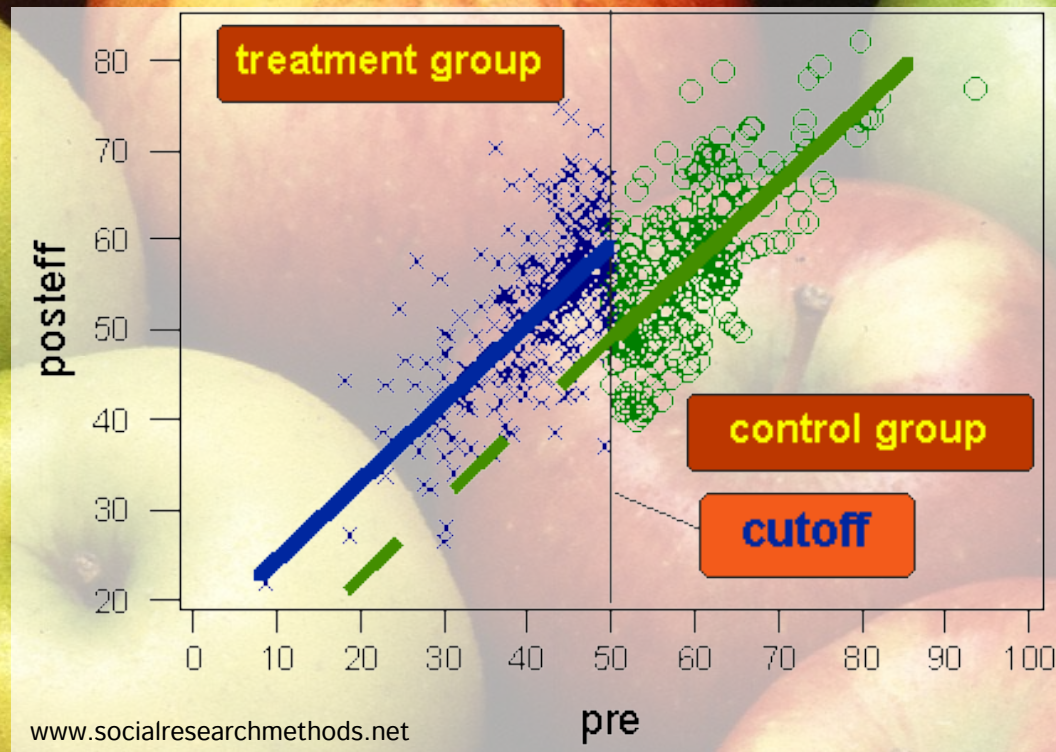
Selección-Historia
Selección-Maduración
Selección-Testeo
Selección-Instrumentación
Selección-Mortalidad
Selección-Regresión (a la media)

Diseño de regresión discontinua

Asignación basada en
umbral de puntuación
en medida antes del
programa

C	O	X	O
C	O		O

Diseño de regresión discontinua



Diseño de regresión discontinua

Necesitan 2.75 veces el número de participantes
de un experimento aleatorio
(para el mismo nivel de precisión estadística)

Otros diseños cuasi-experimentales

Proxy pretest

N	O ₁	X	O ₂
N	O ₁		O ₂

NO RECOMENDADO

Otros diseños cuasi-experimentales

Muestras separadas
pre-post

N_1	0		
N_1		X	0
N_2	0		
N_2			0

No permite emparejar respuestas pre y post.
En general, sólo permite comparar medias.

Otros diseños cuasi-experimentales

N	R ₁	O ₁	O ₂		
	R ₁			X	O
N	R ₂	O ₁	O ₂		
	R ₂				O

Doble pre-test

N	R ₁	O			
	R ₁		X	O	
N	R ₂	O			
	R ₂			O	

N		O	O	X	O
N		O	O		O

No susceptible a selección-madurez

Otros diseños cuasi-experimentales

Replicaciones
alternadas

N	O	X	O		O
N	O		O	X	O

Doble implementación

Otros diseños cuasi-experimentales

Variables dependientes
no equivalentes

N	O ₁	X	O ₁
	O ₂		O ₂

Desplazamiento del
punto de regresión

N _(n=1)	O	X	O
N	O		O

Reduciendo amenazas a la validez



Reduciendo amenazas a la validez

1. Argumentación



Reduciendo amenazas a la validez

2. Medidas/observaciones



Reduciendo amenazas a la validez

3. Diseño

Evitar explicaciones alternativas (p.ej., añadir grupos, medidas...)



Reduciendo amenazas a la validez

4. Análisis

Ajustes (p.ej., abandono)



Reduciendo amenazas a la validez

5. Acciones preventivas





Otras consideraciones acerca del diseño



1. Expandir en el tiempo

(Establece línea base)



2. Expandir programas con diferentes niveles



3. Expandir observaciones

(útil si no hay pretest; ej. memoria, recuerdo)



4. Expandir grupos



Reglas del buen diseño



1. Fuerte base teórica



2. Reflejar el contexto

(competición, interacciones...)



3. Realizable



4. Redundante



5. Eficiente



Conclusión

Análisis

Diseño

Medida

Muestra

Problema

Análisis

1. Preparación de los datos
2. Descripción de los datos (estadística descriptiva)
3. Contraste de hipótesis y modelos (estadística inferencial)



Conclusión

Análisis

Diseño

Medida

Muestra

Problema

Validez de las conclusiones

The End

Posibles amenazas

1. No hay relación pero concluimos que sí
2. Hay relación pero concluimos que no

¿Por qué podemos errar?

1. Violar las suposiciones de los test estadísticos
2. “Ir de pesca”
3. Fiabilidad de las medidas
4. Fiabilidad del programa
5. Heterogeneidad de los participantes

Potencia estadística



*Probabilidad de que la hipótesis nula sea rechazada
cuando la hipótesis alternativa es verdadera*

Potencia estadística



Probabilidad de que la hipótesis nula sea rechazada cuando la hipótesis alternativa es verdadera

Componentes

1. Tamaño de la muestra
2. Tamaño del efecto
3. Nivel de significación

Potencia estadística



	H_0 verdadera, H_1 falsa. En realidad... <ul style="list-style-type: none"> • No existe relación • No hay diferencia • Nuestra teoría es errónea 	H_0 falsa, H_1 verdadera. En realidad... <ul style="list-style-type: none"> • Sí hay relación • Sí hay diferencia • Nuestra teoría es correcta
<p>Aceptamos H_0, rechazamos H_1</p> <p>Decimos...</p> <ul style="list-style-type: none"> • “No hay relación” • “No hay diferencias” • “Nuestra teoría es errónea” 	<p>$1-\alpha$ (e.g., .95) NIVEL DE CONFIANZA</p> <p>Las probabilidades de decir que no hay relación o diferencias cuando en realidad no existen. Las probabilidades de correctamente no confirmar nuestra teoría. <i>El 95% de las veces, cuando no hay efecto, diremos que éste no existe</i></p>	<p>β (e.g., .20) ERROR DE TIPO II</p> <p>Las probabilidades de decir que no hay relación o diferencias cuando en realidad sí existen. Las probabilidades de no confirmar nuestra teoría cuando es cierta. <i>El 20% de las veces, cuando hay un efecto diremos que no existe</i></p>
<p>Rechazamos H_0, aceptamos H_1</p> <p>Decimos...</p> <ul style="list-style-type: none"> • “Sí hay relación” • “Sí hay diferencia” • “Nuestra teoría es correcta” 	<p>α (e.g., .05) ERROR DE TIPO I (NIVEL DE SIGNIFICACIÓN)</p> <p>Las probabilidades de decir que sí hay relación o diferencias cuando en realidad no existen. Las probabilidades de confirmar incorrectamente nuestra teoría. <i>El 5% de las veces, cuando no hay efecto nosotros diremos que sí</i></p>	<p>$1-\beta$ (e.g., .80) POTENCIA</p> <p>Las probabilidades de decir que sí hay relación o diferencias cuando sí existen. Las probabilidades de confirmar que nuestra teoría es correcta. <i>En el 80% de los casos, cuando decimos que existe un efecto, es que éste existe</i></p>

Inferencia estadística



Saeed

Imagen: www.mundiario.com





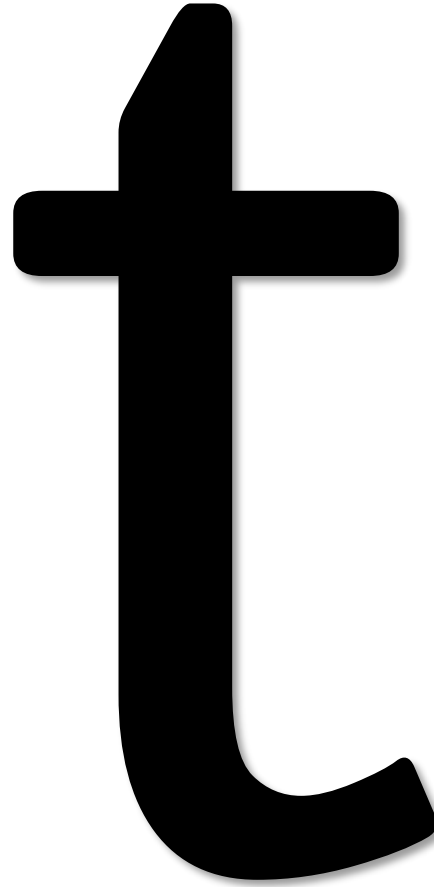
GLM

t

P

Mayor potencia estadística
Estimaciones más precisas

**Quando se cumplen
las suposiciones**



NP

Más robustas

Student

Welch

Mann-Whitney

Wilcoxon

Análisis más habituales

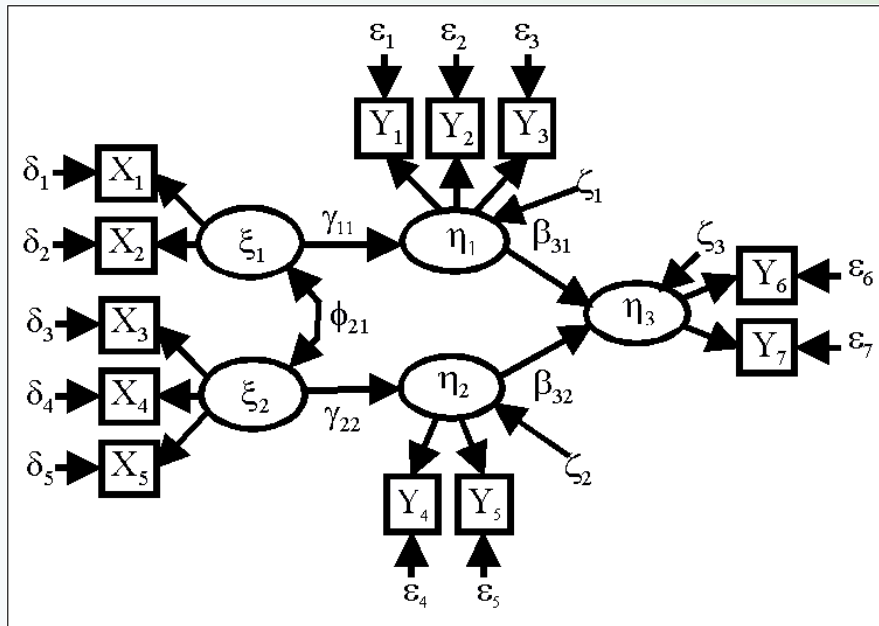
	Diseño	Técnica
Análisis experimental	Experimento aleatorizado 2 grupos sólo posttest	t-test, ANOVA
	Experimento factorial	ANOVA
	Bloques aleatorizados	ANOVA (con bloqueo)
	Covarianzas	ANCOVA
Análisis cuasi-experimental	Grupos no equivalentes	ANCOVA con corrección de fiabilidad
	Regresión discontinua	Regresión polinómica

... y variantes NP

¿Y las regresiones?



Modelos más complejos



SEM
(PLS-PM)

Reglas selección técnicas análisis multivariante (Hair et al., 2010)

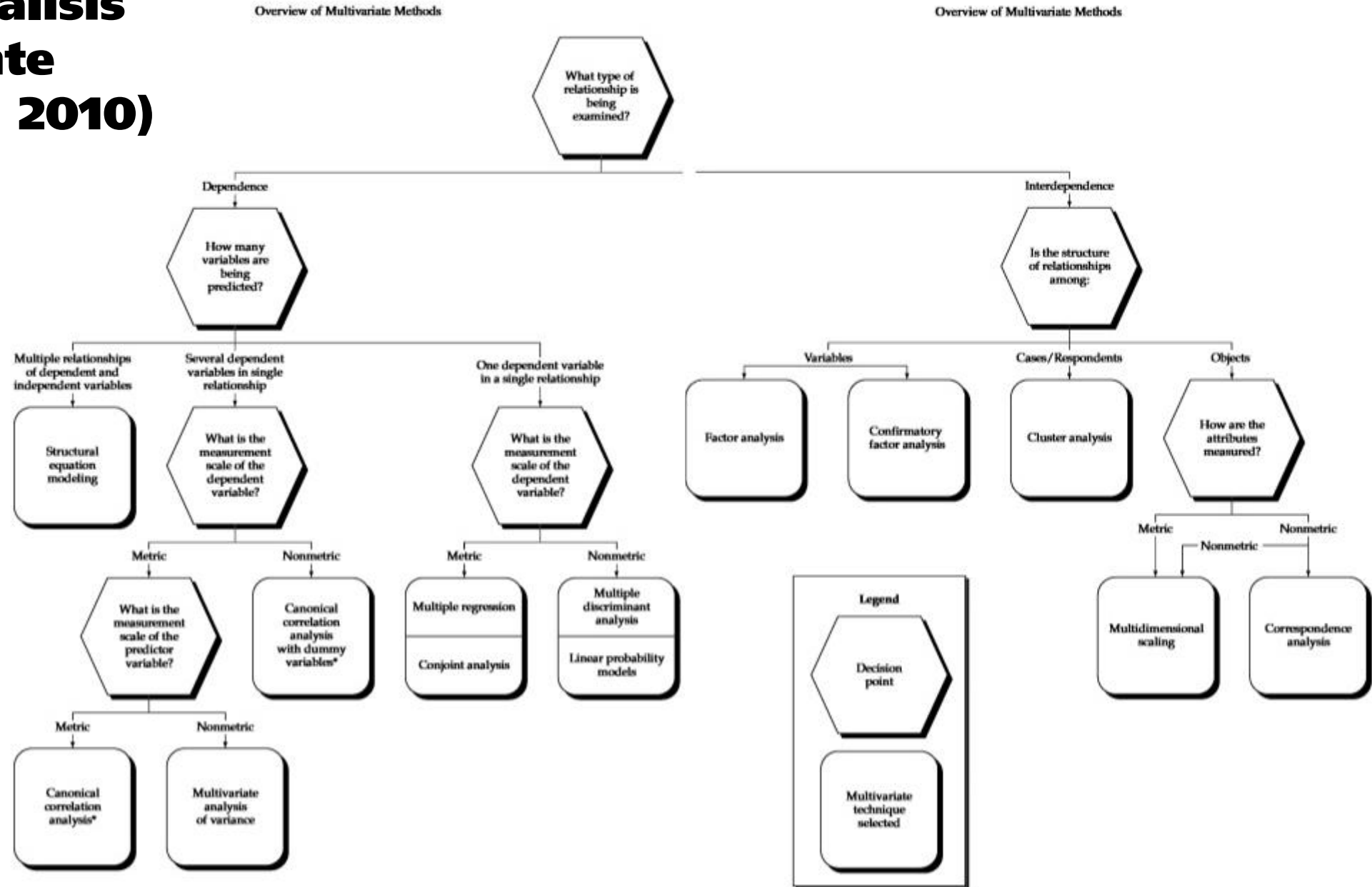
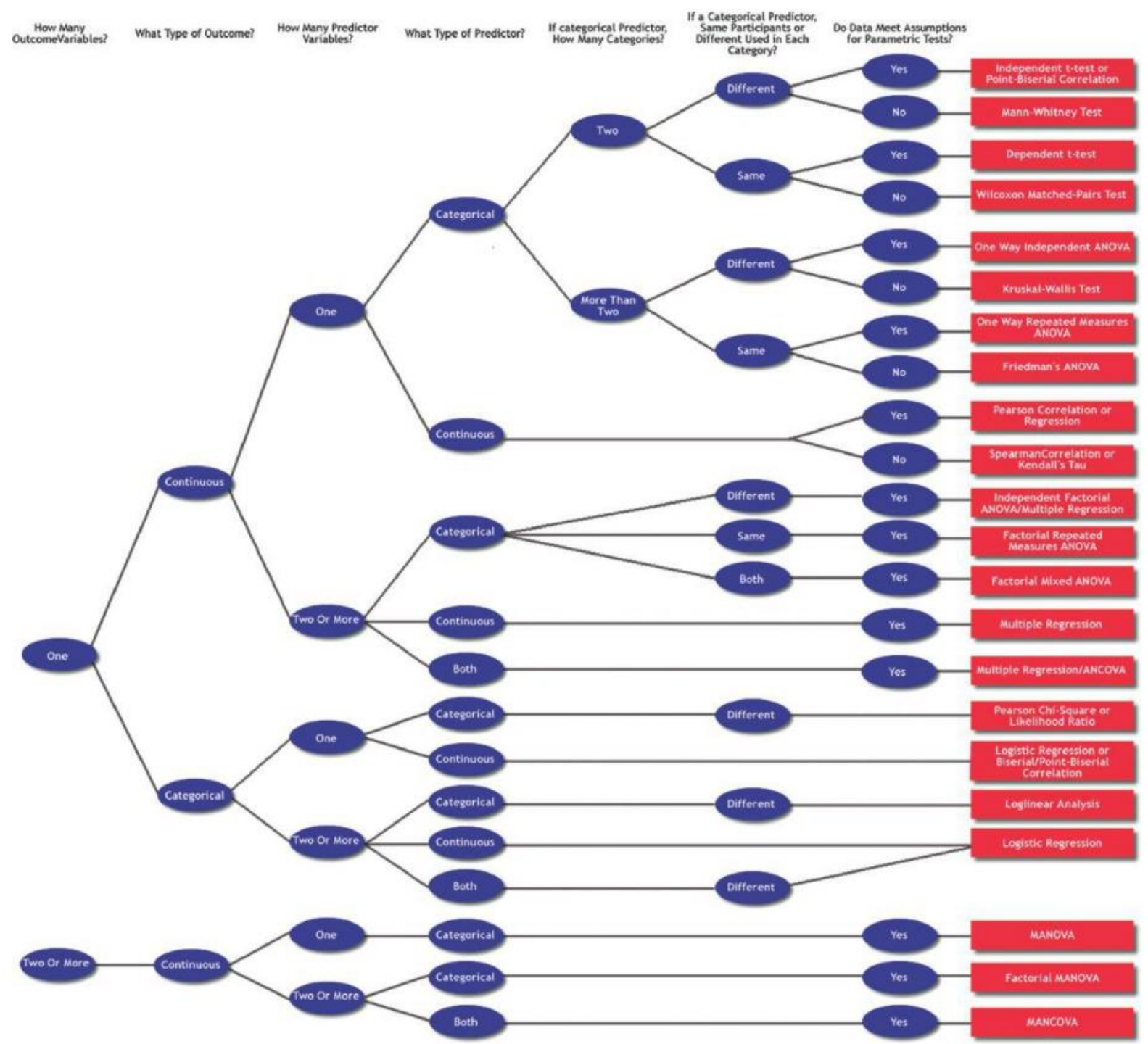


FIGURE 1 Selecting a Multivariate Technique

Reglas selección técnicas análisis (Field, 2013)





That's all Folks!